

TITLE OF THE INVENTION

Coding Device and Coding Method

BACKGROUND OF THE INVENTION

Field of the Invention

5 The present invention relates to a coding device and a coding method for coding a video signal by using a video packet having a length limit set thereto which is related to a portable telephone, a TV telephone system and the like, for example.

Description of the Background Art

10 Fig. 6 is a block diagram showing a conventional coding device described in “Everything about MPEG – 4” (Institute of Industrial Research) P. 39 to P. 40, for example, Fig. 7 is a diagram illustrating an input signal of the conventional coding device, Figs. 8A to 8D are diagrams illustrating a structure of a bit stream, and Fig. 9 is a diagram illustrating a position (arrangement) of a video packet over a screen (display state).

15 In Fig. 6, the reference numeral 1 denotes a subtracter for receiving an external input signal (a luminance signal, a color difference signal or the like) sent externally as a first input. An output of the subtracter 1 is input to a DC / AC predictor 4 for predicting a quantized value of each component of a direct current (DC) and an alternating current (AC) and a reverse quantizer 6 through DCT (Discrete Cosine Transform) means 2 and a quantizer 3. Moreover, an output of the DC / AC predictor 4 is sent to a first input of
20 variable – length coding means 5, and the variable – length coding means 5 outputs a bit stream.

 On the other hand, an output of the reverse quantizer 6 to which an output of the quantizer 3 is input is sent to a first input of an adder 8 through reverse DCT means 7. An output of the adder 8 is sent to a memory 9, and an output of the memory 9 is sent to a
25 first input of predicted image forming means 10 and a first input of motion detecting

means 11.

An external input signal is sent to a second input of the motion detecting means 11, and an output of the motion detecting means 11 is sent to a second input of the predicted image forming means 10 and a motion vector predictor 12.

5 An output of the motion vector predictor 12 is sent to a second input of the variable – length coding means 5. Moreover, an output of the predicted image forming means 10 is sent to a second input of the subtracter 1 and a second input of the adder 8.

Next, an operation will be described. First of all, a video signal is divided into macroblocks to be basic processing units as shown in Fig. 7 and is input as an external
10 input signal (the external input signal is basically input as a macroblock, and means for generating a macroblock may be provided in a former stage such that a conversion into a macroblock is carried out even if the macroblock is directly input) .

More specifically, in the case in which a video signal to be input is 4 : 2 : 0 (which indicates that the number of pixels of luminance information Y is a double in
15 horizontal and vertical directions for the number of pixels of color difference information Cb and Cr), a size of 16 pixels x 16 lines of the luminance signal (Y) becomes equal to that of 8 pixels x 8 lines of two color difference signals (Cb, Cr) over a screen.

Accordingly, six blocks of 8 pixels x 8 lines (including four blocks for the luminance signal and two blocks for the color difference signal) constitute one
20 macroblock.

It is premised that a Video Object Plane (VOP which is a unit image) to be input as an external input has a rectangular shape and is identical to a frame.

Each block is subjected to the discrete cosine transform (DCT) and is quantized in the quantizer 3. After a coefficient of each component of the DC and the AC is
25 predicted in the DC / AC predictor 4, a DCT coefficient thus quantized is variable –

length coded together with additional information such as a quantization parameter.

The foregoing implies intracoding (which is also referred to as in – frame coding). A VOP applying the intracoding to all the macroblocks is referred to as an I – VOP (Intra – VOP).

5 On the other hand, the quantized DCT coefficient is reversely quantized in the reverse quantizer 6 and is decoded by the reverse DCT in the reverse DCT means 7, and a decoded image is stored in the memory 9 through the adder 8. The decoded image stored in the memory 9 is used when intercoding (which is also referred to as interframe coding) is to be carried out.

10 In the case of the intercoding, a motion vector indicative of a motion of a macroblock which is input as an external input signal is detected in the motion detecting means 11. The motion vector indicates such a position that an error is minimized with respect to the input macroblock in the decoded images stored in the memory 9.

 The predicted image forming means 10 forms a predicted image based on the
15 motion vector detected by the motion detecting means 11.

 Subsequently, a differential signal between the input macroblock and the predicted image formed by the predicted image forming means 10 is obtained, is subjected to the DCT in the DCT means 2 and is quantized in the quantizer 3.

 A transformation coefficient thus quantized is variable – length coded
20 (inter-coded) together with the motion vector thus predicted and coded and additional information such as a quantization parameter. Moreover, the quantized DCT coefficient is reversely quantized in the reverse quantizer 6 and is subjected to the reverse DCT in the reverse DCT means 7, and is then added to the predicted image by the adder 8 and is stored in the memory 9.

25 The intercoding includes one – way prediction in which prediction is carried out

based on only a former VOP on a time basis in order of display of the image and
 bidirectional prediction in which prediction is carried out based on former and latter
 VOPs on a time basis. The VOP coded through the one – way prediction will be
 referred to as a P – VOP (Predictive VOP) and the VOP coded through the bi-directional
 5 prediction will be referred to as a B – VOP (Bidirectionally Predictive VOP).

Next, a structure of a bit stream output from the variable – length coding means
 5 will be described with reference to Figs. 8A to 8D. As shown in Fig. 8A, a bit stream
 of 1VOP is constituted by (a bit stream of) one video packet or more.

One video packet is formed by coded data of one macroblock or more. For a
 10 first video packet of the VOP, a VOP header is attached to a head and a stuff bit for a byte
 alignment is attached to an end (Fig. 8B).

In the case of second and succeeding video packets, Resync Marker for
 detecting a head of the video packet and a video packet header are attached to a head, and
 a stuff bit is attached to an end (Fig. 8C).

15 The stuff bit is added up to a termination (break) of the video packet in a unit of
 1 to 8 bits in order to adjust the byte alignment to be attached to the end of the video
 packet and the meaning thereof is distinguished from that of a stuffing which will be
 described below.

As shown in Fig. 8D, moreover, an optional number of stuffings can also be put
 20 in the video packet. For example, in the case of MPEG4 Video, the stuffing is referred
 to as a stuffing macroblock and can be put in an optional video packet in the same manner
 as the macroblock. The stuffing is discarded (is not substantially utilized) on the
 decoder side.

The stuffing is used as a word having 9 bits or 10 bits for the stuffing
 25 irrespective of the byte alignment (for example, the termination of the video packet is

adjusted) and is inserted between the macroblocks, of which meaning is distinguished from the meaning of the stuff bit.

An optional number of macroblocks can be put in one video packet. In the case in which error propagation is taken into consideration, it is generally preferable that a code volume of each video packet should be almost constant. In the case in which the code volume of the video packet is thus set to be almost constant, a rate (area) occupied by each video packet in the 1VOP is not constant as shown in Fig. 9.

In the conventional coding device described above, there has not been considered control of the code volume which is to be carried out when a length of the video packet is limited.

For example, in the case in which a reversible variable – length code is to be used in the variable – length coding means 5, the decoder decodes the variable – length code in a reverse direction from an end of the video packet even if an error is made in an operation for decoding the variable – length code in a forward direction from a head of the video packet. Thus, the variable – length code can be decoded.

In this case, it is necessary to retain one video packet in a receiving buffer on the decoder side. Therefore, a limit is sometimes set to a length of the video packet in order to define a size of the receiving buffer.

In such a case, a coding device should control a code volume such that the length of each video packet is set to be a predetermined length or less.

Moreover, the coding device should manage a volume of generated codes such that a transmitting buffer (not shown) which is provided in a latter stage of the variable – length coding means 5 does not cause an overflow and an underflow.

The quantization parameter to be used in the quantizer 3 is usually adjusted to increase or decrease the code volume. If the code volume is extremely small as in a

static image, it is necessary to insert the stuffing, thereby increasing the code volume such that the transmitting buffer does not cause the underflow.

The stuffing does not have information which is substantially related to the decoding. Therefore, it is desirable that the stuffing should not be inserted if possible.

- 5 For this reason, generally, a minimum stuffing is inserted if the code volume is small after the 1VOP is completely coded.

In the case in which the limit is set to the length of the video packet, the stuffing cannot perfectly enter one video packet when the stuffing is inserted after the 1VOP is completely coded.

- 10 For example, in the case of a static image formed by computer graphics, few codes are generated if the coding is carried out with the P – VOP. On the other hand, in such a structure that the static image is to be coded, a signal indicative of the underflow is output from the transmitting buffer and an operation is carried out to insert the stuffing based on the signal.

- 15 When the stuffing is inserted into a last video packet of the VOP according to the operation, it is sometimes generated (inserted) beyond the limit of the length of the video packet. On condition that a limit is set to a capacity per video packet and the video packet having only the stuffing is prohibited, there has conventionally been a problem in that the length limit of the video packet cannot be maintained or the video
20 packet having only the stuffing is generated.

SUMMARY OF THE INVENTION

- A first aspect of the present invention is directed to a coding device comprising coding means for coding an external input signal in a macroblock unit, first storing means for storing a code output from the coding means, second storing means for storing an
25 output from the first storing means, and code volume control means for controlling

transfer of the code stored in the first storing means to the second storing means based on a code volume of the code obtained by the coding means such that a length of a video packet constituted by the code is a predetermined length or less.

A second aspect of the present invention is directed to the coding device,
 5 wherein the code volume control means controls storage of a stuffing in the second storing means based on a minimum code volume obtained for each unit image constituted by a video packet which is required for coding the unit image.

A third aspect of the present invention is directed to the coding device, wherein the code volume control means determines a minimum code volume T_{min} to satisfy a
 10 following equation:

$$T_{min} \geq 2 \cdot R_p - B$$

$$R_p = R / F$$

wherein a bit count read from the second storing means in a unit image is represented by R_p , an occupancy in the second storing means (a data capacity stored in the second
 15 storing means) is represented by B , a bit rate read from the second storing means is represented by R , and a rate of a unit image to be coded is represented by F .

A fourth aspect of the present invention is directed to the coding device, wherein the code volume control means determines a minimum code volume T_{min} to satisfy a following equation:

20
$$T_{min} \geq v_{bv_bits} + 2 \cdot R_p - v_{bv_bs}$$

$$R_p = R / F$$

wherein a bit count read from the second storing means in a unit image is represented by R_p , an occupancy of a VBV buffer in a last unit image (a data capacity retained in the VBV buffer) is represented by v_{bv_bits} , a size of the VBV buffer is represented by
 25 v_{bv_bs} , a bit rate read from the second storing means is represented by R , and a rate of a

unit image to be coded is represented by F.

A fifth aspect of the present invention is directed to the coding device, wherein the code volume control means determines a minimum code volume T_{min} based on a following equation or a value having a result equivalent to a result of the equation:

$$T_{min} = \max (2 \cdot R_p - B, vbv_bits + 2 \cdot R_p - vbv_bs)$$

$$R_p = R / F$$

wherein a bit count read from the second storing means in a unit image is represented by R_p, an occupancy in the second storing means (a data capacity stored in the second storing means) is represented by B, an occupancy of a VBV buffer in a last unit image (a data capacity retained in the VBV buffer) is represented by vbv_bits, a size of the VBV buffer is represented by vbv_bs, a bit rate read from the second storing means is represented by R, and a rate of a unit image to be coded is represented by F.

A sixth aspect of the present invention is directed to the coding device, wherein the bit rate R read from the second storing means is variable.

A seventh aspect of the present invention is directed to the coding device, wherein the code volume control means inserts a stuffing into a video packet until a first relationship is not satisfied, when a present code volume of a unit image including a last coded macroblock constituting the unit image is smaller than the minimum code volume T_{min} of the unit image and a number M of macroblocks to be coded subsequently to the last coded macroblock, a predetermined length VPlen of the video packet, the minimum code volume T_{min} and the present code volume Sc have the first relationship:

$$M \cdot VPlen < T_{min} - Sc,$$

the code volume control means constitutes a video packet next to the video packet by a macroblock next to the last coded macroblock without inserting a stuffing into the video packet, when the first relationship is not established and the number M of macroblocks,

the length VPlen of a video packet, the minimum code volume Tmin and the present code volume Sc have a second relationship:

$$(M - 1) \cdot VPlen < Tmin - Sc.$$

An eighth aspect of the present invention is directed to a coding method comprising the steps of (a) coding an external input signal in a macroblock unit, (b) storing a code obtained at the step (a), (c) controlling an output of the code stored at the step (b) such that a length of a video packet constituted by the code obtained at the step (a) is a predetermined length or less based on a code volume of the code, and (d) storing the output controlled by the step (c).

A ninth aspect of the present invention is directed to the coding method, wherein the step (c) serves to control storage of a stuffing at the step (d) based on a minimum code volume obtained for each unit image constituted by a video packet which is required for coding the unit image.

A tenth aspect of the present invention is directed to the coding method, wherein the step (c) serves to determine a minimum code volume Tmin to satisfy a following equation:

$$Tmin \geq 2 \cdot Rp - B$$

$$Rp = R / F$$

wherein a bit count read by the step (d) in a unit image is represented by Rp, an occupancy in the step (d) (a data capacity stored in the step (d)) is represented by B, a bit rate read by the step (d) is represented by R, and a rate of a unit image to be coded is represented by F.

An eleventh aspect of the present invention is directed to the coding method, wherein the step (c) serves to determine a minimum code volume Tmin to satisfy a following equation:

$$T_{min} \geq vbv_bits + 2 \cdot R_p - vbv_bs$$

$$R_p = R / F$$

wherein a bit count read by the step (d) in a unit image is represented by R_p , an occupancy of a VBV buffer in a last unit image (a data capacity retained in the VBV buffer) is represented by vbv_bits , a size of the VBV buffer is represented by vbv_bs , a
 5 bit rate read by the step (d) is represented by R , and a rate of a unit image to be coded is represented by F .

A twelfth aspect of the present invention is directed to the coding method, wherein the step (c) determines a minimum code volume T_{min} based on a following
 10 equation or a value having a result equivalent to a result of the equation:

$$T_{min} = \max(2 \cdot R_p - B, vbv_bits + 2 \cdot R_p - vbv_bs)$$

$$R_p = R / F$$

wherein a bit count read by the step (d) in a unit image is represented by R_p , an occupancy in the step (d) (a data capacity stored in the step (d)) is represented by B , an
 15 occupancy of a VBV buffer in a last unit image (a data capacity retained in the VBV buffer) is represented by vbv_bits , a size of the VBV buffer is represented by vbv_bs , a bit rate read by the step (d) is represented by R , and a rate of a unit image to be coded is represented by F .

A thirteenth aspect of the present invention is directed to the coding method,
 20 wherein the bit rate R at which a code stored at the step (d) is read is variable.

A fourteenth aspect of the present invention is directed to the coding method, wherein the step (c) serves to insert a stuffing into a video packet until a first relationship is not satisfied, when a present code volume of a unit image including a last coded macroblock constituting the unit image is smaller than the minimum code volume T_{min}
 25 of the unit image and a number M of macroblocks to be coded subsequently to the last

coded macroblock, a predetermined length VPlen of the video packet, the minimum code volume Tmin and a present code volume Sc have a first relationship: $M \times VPlen < Tmin - Sc$, the code volume controlling step serves to constitute a video packet next to the video packet by a macroblock next to the last coded macroblock without inserting a stuffing into the video packet, when the first relationship is not established and the number M of macroblocks, the length VPlen of a video packet, the minimum code volume Tmin and the present code volume Sc have a second relationship: $(M - 1) \times VPlen < Tmin - Sc$.

According to the present invention, the above-mentioned structure can give the following effects.

According to the first aspect of the present invention, the coding device comprises coding means for coding an external input signal in a macroblock unit, first storing means for storing a code output from the coding means, second storing means for storing an output from the first storing means, and code volume control means for controlling transfer of the code stored in the first storing means to the second storing means based on a code volume of the code obtained by the coding means such that a length of a video packet constituted by the code is a predetermined length or less. Also in the case in which the video packet has a length limit, therefore, a structure thereof can be obtained corresponding to the limit.

According to the second aspect of the present invention, the code volume control means in the coding device controls storage of a stuffing in the second storing means based on a minimum code volume obtained for each unit image required for coding the unit image constituted by a video packet. Also in the case of an image having a small generated code volume such as a static image, therefore, it is possible to insert a minimum stuffing.

According to the third aspect of the present invention, the code volume control means in the coding device calculates a minimum code volume T_{min} based on the following equation:

$$T_{min} \geq 2 \cdot R_p - B$$

$$R_p = R / F$$

wherein T_{min} indicates a minimum code volume, R_p indicates a bit count read from the second storing means in a unit image, R indicates a bit rate read from the second storing means, F indicates a rate of a unit image to be coded, and B indicates an occupancy in the second storing means. Therefore, an underflow of the second storing means can be prevented.

According to the fourth aspect of the present invention, the code volume control means in the coding device calculates a minimum code volume T_{min} based on the following equation:

$$T_{min} \geq vbv_bits + 2 \cdot R_p - vbv_bs$$

$$R_p = R / F$$

wherein T_{min} indicates a minimum code volume, R_p indicates a bit count read from the second storing means in a unit image, R indicates a bit rate read from the second storing means, F indicates a rate of a unit image to be coded, vbv_bits indicates an occupancy of a VBV buffer in a last unit image, and vbv_bs indicates a size of the VBV buffer. Therefore, an overflow of the VBV buffer can be prevented.

According to the fifth aspect of the present invention, the code volume control means in the coding device calculates a minimum code volume T_{min} based on the following equation or a value having a result equivalent to a result of the equation:

$$T_{min} = \max(2 \cdot R_p - B, vbv_bits + 2 \cdot R_p - vbv_bs)$$

$$R_p = R / F$$

wherein T_{min} indicates a minimum code volume, R_p indicates a bit count read from the second storing means in a unit image, R indicates a bit rate read from the second storing means, F indicates a rate of a unit image to be coded, B indicates an occupancy in the second storing means, vbv_bits indicates an occupancy of a VBV buffer in a last unit image, and vbv_bs indicates a size of the VBV buffer. Therefore, both an underflow of the second storing means and an overflow of the VBV buffer can be avoided.

According to the sixth aspect of the present invention, the bit rate R read from the second storing means is variable. Therefore, the underflow of the second storing means or the overflow of the VBV buffer can be avoided effectively.

According to the seventh aspect of the present invention, when a present code volume of a unit image including a last coded macroblock constituting the unit image is smaller than the minimum code volume T_{min} of the unit image and a number M of macroblocks to be coded subsequently to the last coded macroblock, a predetermined length $VPlen$ of a video packet, the minimum code volume T_{min} and a present code volume Sc have a first relationship: $M \cdot VPlen < T_{min} - Sc$, the code volume control means in the coding device inserts a stuffing into the video packet until the first relationship is not satisfied, and

when the first relationship is not established and the number M of macroblocks, the length $VPlen$ of a video packet, the minimum code volume T_{min} and the present code volume Sc have a second relationship: $(M - 1) \cdot VPlen < T_{min} - Sc$, the code volume control means constitutes a video packet next to the video packet by a macroblock next to the last coded macroblock without inserting a stuffing into the video packet. Therefore, a video packet having only a stuffing can be prevented from being generated, and an underflow of a transmitting buffer or an overflow of the VBV buffer can be prevented by inserting a minimum stuffing.

According to the eighth aspect of the present invention, the coding method comprises the steps of (a) coding an external input signal in a macroblock unit, (b) storing a code obtained at the coding step, (c) controlling an output of the code stored at the step (b) such that a length of a video packet constituted by the code is a predetermined length or less based on a code volume of the code obtained at the step (a) , and (d) storing the output controlled at the step (c). Therefore, also in the case in which the video packet has a length limit, a structure thereof can be obtained corresponding to the limit.

According to the ninth aspect of the present invention, the step (c) in the coding method serves to control storage of a stuffing at the step (d) based on a minimum code volume obtained for each unit image constituted by a video packet which is required for coding the unit image. Also in the case of an image having a small generated code volume such as a static image, therefore, it is possible to insert a minimum stuffing.

According to the tenth aspect of the present invention, the step (c) in the coding method calculates a minimum code volume T_{min} based on the following equation:

$$T_{min} \geq 2 \cdot R_p - B$$

$$R_p = R / F$$

wherein T_{min} indicates a minimum code volume, R_p indicates a bit count with which a code stored at the step (d) in a unit image is read, R indicates a bit rate at which a code stored at the step (d) is read, F indicates a rate of a unit image to be coded, and B indicates an occupancy in the storage of a code at the step (d).

According to the eleventh aspect of the present invention, the step (c) in the coding method calculates a minimum code volume T_{min} based on the following equation:

$$T_{min} \geq v_{bv_bits} + 2 \cdot R_p - v_{bv_bs}$$

$$R_p = R / F$$

wherein Tmin indicates a minimum code volume, Rp indicates a bit count with which a code stored at the step (d) in a unit image is read, R indicates a bit rate at which a code stored at the step (d) is read, F indicates a rate of a unit image to be coded, vbv_bits indicates an occupancy of a VBV buffer in a last unit image, and vbv_bs indicates a size of the VBV buffer. Therefore, an overflow of the VBV buffer can be prevented.

According to the twelfth aspect of the present invention, the step (c) in the coding method serves to calculate a minimum code volume Tmin based on the following equation or a value having a result equivalent to a result of the equation:

$$Tmin = \max (2 \cdot Rp - B, vbv_bits + 2 \cdot Rp - vbv_bs)$$

$$Rp = R / F$$

wherein Tmin indicates a minimum code volume, Rp indicates a bit count with which a code stored at the step (d) in a unit image is read, R indicates a bit rate at which a code stored at the step (d) is read, F indicates a rate of a unit image to be coded, B indicates an occupancy in the storage of a code at the step (d), vbv_bits indicates an occupancy of a VBV buffer in a last unit image, and vbv_bs indicates a size of the VBV buffer. Therefore, both an underflow of the second storing means and an overflow of the VBV buffer can be avoided.

According to the thirteenth aspect of the present invention, the bit rate R at which a code stored at the step (d) is read is variable. Therefore, the underflow of the second storing means or the overflow of the VBV buffer can be avoided effectively.

According to the fourteenth aspect of the present invention, when a code volume of a unit image including a last coded macroblock constituting the unit image is smaller than the minimum code volume Tmin of the unit image and a number M of macroblocks to be coded subsequently to the last coded macroblock, a predetermined length VPlen of a video packet, the minimum code volume Tmin and a present code

volume Sc have a first relationship: $M \cdot VPlen < Tmin - Sc$,

the code volume controlling step in the coding method serves to insert a stuffing into the video packet until the first relationship is not satisfied, and

when the first relationship is not established and the number M of macroblocks,
 5 the length $VPlen$ of a video packet, the minimum code volume $Tmin$ and the present code volume Sc have a second relationship: $(M - 1) \cdot VPlen < Tmin - Sc$,
 the step (c) serves to constitute a video packet next to a video packet by a macroblock next to the coded macroblock without inserting a stuffing into the video packet.
 Therefore, a video packet having only a stuffing can be prevented from being generated,
 10 and an underflow of a transmitting buffer or an overflow of the VBV buffer can be prevented by inserting a minimum stuffing.

In order to solve the above-mentioned problems, it is an object of the present invention to provide a coding device which does not generate a video packet having only a stuffing but can insert a minimum stuffing to satisfy a length limit of a video packet in
 15 the case in which the limit is set to the length of the video packet.

These and other objects, features, aspects and advantages of the present invention will become more apparent from the following detailed description of the present invention when taken in conjunction with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

20 Fig. 1 is a block diagram showing a coding device according to a first embodiment,

Fig. 2 is a diagram illustrating a state of a temporary buffer and a transmitting buffer according to the first embodiment (in the case of an I - VOP),

Fig. 3 is a diagram illustrating a state of the temporary buffer and the
 25 transmitting buffer according to the first embodiment (in the case of a P - VOP),

Fig. 4 is a flow chart for explaining an operation in the coding device according to the first embodiment,

Figs. 5A to 5C are diagrams showing a structure of a video packet according to the first embodiment,

5 Fig. 6 is a block diagram showing a conventional coding device,

Fig. 7 is a diagram showing an external input signal to be sent to the conventional coding device,

Figs. 8A to 8D are diagrams showing a structure of a bit stream in the conventional coding device, and

10 Fig. 9 is a diagram showing a position of a video packet over a screen in the conventional coding device.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention will be described specifically with reference to the drawings showing embodiments.

15 First Embodiment

Fig. 1 shows a coding device according to a first embodiment of the present invention. In Fig. 1, the reference numeral 1 denotes a subtracter for receiving an external input signal as a first input. An output of the subtracter 1 is input to a DC/AC predictor 4 and a reverse quantizer 6 through DCT means 2 and a quantizer 3. An
20 output of the DC/AC predictor 4 is sent to a first input of variable – length coding means 5a.

On the other hand, an output of the reverse quantizer 6 is sent to a first input of an adder 8 through reverse DCT means 7. An output of the adder 8 is sent to a memory 9, and an output of the memory 9 is sent to a first input of predicted image forming means
25 10 and a first input of motion detecting means 11.

The external input signal is sent to a second input of the motion detecting means 11, and an output of the motion detecting means 11 is sent to a second input of the predicted image forming means 10 and a motion vector predictor 12. An output of the predicted image forming means 10 is sent to a second input of the subtracter 1 and a
 5 second input of the adder 8.

Moreover, an output of the motion vector predictor 12 is sent to a second input of the variable – length coding means 5a. Coding means is constituted by inclusion from the subtracter 1 for inputting the external input signal to the variable – length coding means 5a for outputting a variable – length code corresponding to the external input
 10 signal (Of course, the structure is only illustrated as an example and a well – known structure capable of carrying out coding corresponding to the external input signal can be used).

A first output of the variable – length coding means 5a is sent to a first input of a temporary buffer 101 (first storing means), and a second output of the variable – length
 15 coding means 5a is sent to an input of code volume control means 102.

A first output of the code volume control means 102 is sent to a second input of the temporary buffer 101, and an output of the temporary buffer 101 is sent to a first input of a transmitting buffer 103 (second storing means). A second output of the code
 20 volume control means 102 is sent to a second input of the transmitting buffer 103, and an output of the transmitting buffer 103 is output (transmitted) as a bit stream.

The bit stream thus output (transmitted) is received on the decoder side and is subjected to a decoding processing.

Next, an operation will be described.

First of all, a video signal is divided into macroblocks to be basic processing
 25 units as shown in Fig. 7 and is then input. For example, in the case in which the video

signal to be input is 4 : 2 : 0, a size of 16 pixels x 16 lines of a luminance signal (Y) is equal to that of 8 pixels x 8 lines of two color difference signals (Cb, Cr) over a screen. Therefore, one macroblock is constituted by six blocks, each of the blocks having 8 pixels x 8 lines.

5 In the case in which intracoding is to be carried out, each block is subjected to DCT and is then quantized. A DCT coefficient thus quantized is predicted by the DC/AC predictor 4 and is then variable – length coded together with additional information such as a quantization parameter. The DCT coefficient thus quantized is decoded through reverse quantization and reverse DCT, and a decoded image is stored in
10 the memory 9.

 In the case in which intercoding is to be carried out, the motion detecting means
11 detects a motion vector indicative of a motion of the input macroblock. The motion vector indicates such a position that the smallest error is made with respect to the input macroblock in the decoded image stored in the memory 9.

15 Based on the motion vector, the predicted image forming means 10 forms a predicted image. Next, a difference between the input macroblock and the predicted image is obtained so that a difference signal is subjected to the DCT and is quantized.

 The DCT coefficient thus quantized is variable – length coded together with the motion vector thus predicted and coded and the additional information such as the
20 quantization parameter. Moreover, after the quantized DCT coefficient is subjected to the reverse quantization and the reverse DCT, it is added to the predicted image and is stored in the memory 9.

 Next, an operation of the variable – length coding means 5a will be described in detail.

25 The variable – length coding means 5a codes the quantized DCT coefficient and

the additional information for each macroblock (a coding step), writes them to the temporary buffer 101 (a first storing step) and outputs a code volume thereof to the code volume control means 102.

In the case of an I – VOP of an MPEG4, for example, an AC component of the quantized DCT coefficient of each block is first one – dimensionally scanned by zigzag scan or the like, and run – length coding for coding a combination of the number of 0s and a coefficient of non – zero is carried out. Coefficient data of each block which are run – length coded are written to the temporary buffer 101.

As shown in Fig. 2, mcbpc obtained by collectively coding MTYPE indicative of a macroblock type and CBPC indicating whether each block for a color difference has a non – zero AC coefficient, dquant indicative of a quantization parameter, a DC component of a DCT coefficient of each block, ac_pred_flag indicating whether AC prediction is carried out, and cbpy indicating whether each block of Y has a non-zero AC coefficient are coded in order after the coefficient data of each block which are stored in the temporary buffer 101, and are written to the temporary buffer 101.

A total of the code volumes is output to the code volume control means 102 for each macroblock.

In the case of a P – VOP of the MPEG4, similarly, the data coded in order shown in Fig. 3 are written to the temporary buffer 101.

The code volume control means 102 collects macroblocks such that a length of each video packet has a predetermined value (VPlen) or less based on a code volume of each macroblock which is output from the variable – length coding means 5a (a code volume controlling step), and transfers them from the temporary buffer 101 to the transmitting buffer 103 (a second storing step).

In the case of the MPEG4, for example, a header is added to a head of the video

packet and is then rearranged and transferred to the transmitting buffer 103 in order of the defined bit stream as shown in Figs. 2 and 3.

Moreover, the code volume control means 102 sets a minimum code volume T_{min} for each VOP such that the transmitting buffer 103 does not cause an underflow or a VBV (Video Buffering Verifier) buffer (a virtual buffer required for receiving a video packet on the receiving side (a required capacity is described in a header of a transmitting bit stream, for example). At least a capacity for the I – VOP is usually set.) does not cause an overflow, and writes a stuffing to the transmitting buffer 103 if necessary and determines a break of the video packet such that the code volume of the VOP is not smaller than the T_{min} .

More specifically, the minimum code volume T_{min} implies a minimum code volume which is required such that the transmitting buffer 103 does not cause the underflow and the VBV buffer does not cause the overflow.

The details of the operation will be described below.

The code volume control means 102 calculates the minimum code volume T_{min} required for the VOP before the coding for each VOP is started. For example, a bit count R_p read from the transmitting buffer 103 for a 1VOP period ($(1 / F)$ sec) is obtained as follows:

$$R_p = R / F$$

wherein a current occupancy of the transmitting buffer 103 in the coding device (a data capacity retained in the transmitting buffer 103) is represented by B (bits. An occupancy in the second storing means), a read bit rate of the transmitting buffer 103 is represented by R (bits / sec), and a rate of a VOP to be coded is represented by F (1 / sec). In order for the transmitting buffer 103 to cause no underflow, therefore, it is sufficient that the occupancy of the transmitting buffer 103 is always R_p or more. Accordingly, it is

preferable that the minimum code volume T_{min} should be set as follows:

$$T_{min} \geq 2 \cdot R_p - B.$$

Moreover, in the case in which the VBV buffer is to be managed, it is sufficient that an occupancy of the VBV buffer is $vbv_bs - R_p$ or less such that the VBV buffer does not cause the overflow, wherein the occupancy of the VBV buffer (a data occupancy retained in the VBV buffer) for a time required for decoding a VOP prior to a current VOP is represented by vbv_bits (the occupancy of the VBV buffer) and a size of the VBV buffer is represented by vbv_bs .

Accordingly, it is preferable that a minimum code volume T_{min} of the current VOP should be set as follows:

$$T_{min} \geq vbv_bits + 2 \cdot R_p - vbv_bs.$$

Since the occupancy vbv_bits of the VBV buffer presumes an occupancy on the receiving side, and is calculated based on a read bit rate of the transmitting buffer 103, for example, it is changed with the passage of time.

Accordingly, the code volume control means 102 sets a minimum code volume T_{min} required for the VOP before the coding of each VOP is started:

$$T_{min} = \max (2 \cdot R_p - B, \\ vbv_bits + 2 \cdot R_p - vbv_bs)$$

($\max(a, b)$ indicates that a or b which is greater is selected as a value).

In the case in which the transmitting buffer 103 of the coding device is empty, it is not necessary to manage the underflow of the transmitting buffer 103 in such a structure that the reading operation of the transmitting buffer 103 is stopped (halted). Therefore, it is preferable that the following equation should be set:

$$T_{min} = vbv_bits + 2 \cdot R_p - vbv_bs.$$

As described above, the vbv_bits is changed on a time basis. Therefore, the

value of the minimum code volume T_{min} is also changed on a time basis and is calculated for each VOP.

Next, the code volume control means 102 calculates a present code volume Sc of the current VOP for each macroblock and decides whether a new video packet is to be constituted at a next macroblock or not and whether a stuffing is to be inserted into a current video packet or not in accordance with a flow chart shown in Fig. 4 and a structure of a video packet shown in Fig. 5.

The total number of macroblocks constituting the VOP is represented by A , a macroblock number of a current macroblock (a last coded macroblock) is represented by K ($0 \leq K \leq A - 1$), and the number M of succeeding macroblocks to be coded (the number M of residual macroblocks) is represented by $A - K - 1$ (that is, $M = A - K - 1$).

In the case in which the present code volume Sc of the VOP including the current macroblock is smaller than the minimum code volume T_{min} of the VOP, if (a first relationship)

$$M \cdot VPlen < T_{min} - Sc \quad (1),$$

a stuffing is inserted into a current video packet to constitute a new video packet at a next macroblock until the relationship in the equation (1) is not satisfied.

If the equation (1) is not established but the following equation is established (a second relationship), the stuffing is not inserted into the current video packet but a new video packet is constituted at a next macroblock:

$$(M - 1) \cdot VPlen < T_{min} - Sc \quad (2).$$

In other cases, the macroblocks are collected to constitute a video packet such that a length of each video packet is $VPlen$ or less as described above.

An operation of the flow chart shown in Fig. 4 will be described below.

In the case in which the equation (1) is established, M residual video packets

can be constituted as shown in Fig. 5A if the number of residual macroblocks is M. Therefore, a code having a code volume of $M \cdot VPlen$ can be generated.

In this case, accordingly, an insufficient code volume can be obtained as follows.

$$5 \quad T_{min} - Sc - M \cdot VPlen$$

In the case of the MPEG4, for example, $(T_{min} - Sc - M \cdot VPlen + L - 1) / L$ stuffing macroblocks are inserted into the current video packet, wherein a code length of the stuffing macroblock is represented by L.

Next, in the case in which the equation (1) is not established but the equation
10 (2) is established, a residual generated code volume of $(M - 1) \cdot VPlen$ is obtained at a maximum as shown in Fig. 5C if a next macroblock is inserted into the current video packet. Therefore, if a generated code volume of the next macroblock is 0, $Sc < T_{min}$ is obtained based on the equation (2).

If a new video packet is constituted at a next macroblock, a residual generated
15 code volume of $M \cdot VPlen$ is obtained at a maximum as shown in Fig. 5B.

If the equation (1) is not satisfied, a relationship of $M \cdot VPlen \geq T_{min} - Sc$ is obtained and a code having a code volume of $M \cdot VPlen$ can be generated after the next macroblock. Therefore, the following equation can be obtained.

$$\begin{aligned} & \text{A code volume for all the macroblocks constituting the VOP} \\ 20 \quad & = M \cdot VPlen + Sc \\ & \geq T_{min} \end{aligned}$$

Accordingly, it is not necessary to insert the stuffing into the current video packet.

Similarly, in the case in which the equation (2) is not established, it can be guaranteed that the code volume of $(M - 1) \cdot VPlen$ is generated in a residual video
25 packet even if a next macroblock is inserted into the current video packet as shown in Fig.

5C. Therefore, it is not necessary to insert the stuffing at the present time.

By thus controlling the code volume based on the flow chart of Fig. 4, a structure of the video packet can be determined to insert a minimum stuffing under the restriction that a maximum length of the video packet is VPlen.

5 While the read rate of the transmitting buffer 103 is represented by R in the setting of the Tmin in the above-mentioned embodiment, similarly, the Tmin can be set such that the underflow of the transmitting buffer 103 or the overflow of the VBV buffer is not caused even if the read rate is not fixed but variable.

10 The case in which the read rate of the transmitting buffer 103 is variable is equivalent to the case in which a maximum transmission rate is determined and is assigned depending on a type of information to be transmitted (such as a video or a speech), for example.

Also in this case, the code volume is controlled based on the flow chart of Fig. 4. Consequently, the insertion of the stuffing and the break of the video packet can be
15 determined to carry out control such that a code volume of each VOP is Tmin or more.

The case in which data partition of the MPEG4 (data are constituted for each macroblock shown in Fig. 2 every category of (1) mcbpc, dquant and DC component, (2) ac_pred_flag and cbpy and (3) coefficient data of each block in the transmitting buffer 103) has been taken as an example in the above-mentioned embodiment. Also in the
20 case of no data partition or H.263, a code volume can be controlled with the same structure as that described above if a video packet has a length limit.

Referring to the category for the data partition, a structure of (1) not_coded, mcbpc and motion vector, (2) cbpy and dquant and (3) coefficient data of each block may be stored in the transmitting buffer 103 as shown in Fig. 3, for example. Basically, it is
25 preferable that the coefficient data of each block and additional information related to the

coefficient data should be classified as categories. Moreover, (1) to (3) are not always required for the category and an optional number of categories can be permitted.

Furthermore, it is apparent that the present invention can also be applied to the case in which an input signal is not 4 : 2 : 0 and the case in which a VOP (unit image) is
5 not rectangular (for example, an optional shape which can be taken by an object in a screen).

While the invention has been shown and described in detail, the foregoing description is in all aspects illustrative and not restrictive. It is therefore understood that numerous modifications and variations can be devised without departing from the scope
10 of the invention.